

This Page Is Inserted by IFW Operations  
and is not a part of the Official Record

## BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

## IMAGES ARE BEST AVAILABLE COPY.

As rescanning documents *will not* correct images,  
please do not report the images to the  
Image Problem Mailbox.

**THIS PAGE BLANK (USPTO)**



09/913,159

#11

(11) Publication number:

0 327 180

A2

(12)

## EUROPEAN PATENT APPLICATION

(21) Application number: 89200230.4

(51) Int. Cl.4: A61K 39/21, C07K 15/04,  
C12N 15/00

(22) Date of filing: 02.02.89

(30) Priority: 03.02.88 US 151976

(71) Applicant: MicroGeneSys, Inc.

(43) Date of publication of application:  
09.08.89 Bulletin 89/32

400 Frontage Road  
West Haven Connecticut 06437(US)

(84) Designated Contracting States:  
AT BE CH DE ES FR GB GR IT LI LU NL SE

(72) Inventor: Smith, Gale E.

125 Michael Drive  
Guilford Connecticut 06437(US)

Inventor: Cochran, Mark A.

415 Brooksvale Avenue  
Hamden, Connecticut 06518(US)

Inventor: Volfovitz, Franklin  
123 York Street

New Haven, Connecticut 06511(US)

(74) Representative: Smulders, Theodorus A.H.J.,  
Ir. et al  
Vereenigde Ocroobureaux Nieuwe Parklaan  
107  
NL-2587 BP 's-Gravenhage(NL)

(54) Vaccine containing polypeptides derived from the envelope gene of human immunodeficiency virus type 1.

(57) An Acquired Immunodeficiency Syndrome (AIDS) vaccine containing the Human Immunodeficiency Virus, Type-1 (HIV-1) envelope proteins is produced from cloned HIV-1 envelope genes in a baculovirus-insect cell vector system. The recombinant HIV-1 proteins are purified, assembled into particles and then adsorbed on an aluminum phosphate adjuvant. The resulting adsorbed recombinant HIV-1 virus envelope protein formulation is highly immunogenic in animals and elicits antibodies which bind to the HIV-1 virus envelope and neutralize the infectivity of the virus in in vitro tests.

EP 0 327 180 A2

**VACCINE CONTAINING POLYPEPTIDES DERIVED FROM THE ENVELOPE GENE OF HUMAN IMMUNODEFICIENCY VIRUS TYPE 1**

**BACKGROUND OF THE INVENTION**

The Human Immunodeficiency Virus Type-1 (HIV-1) is a retrovirus which causes a systemic infection with a major pathology in the immune system and is the etiological agent (Barre-Sinoussi et al. 1983; Popovic et al. 1984) responsible for Acquired Immunodeficiency Syndrome (AIDS). Clinical isolates of HIV-1 have also been referred to Lymphadenopathy-Associated Virus (Feorino et al. 1984), Human T-cell Leukemia Virus III (Popovic et al. 1984) and AIDS-related Virus (Levy, et al. 1984).

AIDS has become pandemic and the development of a vaccine has become a major priority for world public health. A high percentage of persons infected with HIV-1 show a progressive loss of immune function due to the depletion of T4 lymphocytes. These T4 cells, as well as certain nerve cells, have a molecule on their surface called CD4. HIV-1 recognizes the CD4 molecule through a receptor located in the envelope of the virus particles enters these cells, and eventually replicates and kills the cell. An effective AIDS vaccine might be expected to elicit antibodies which would bind to the envelope of HIV-1 and prevent it from infecting T4 lymphocytes or other susceptible cells.

Vaccines are generally given to healthy individuals before they are exposed to a disease organism as an immune prophylactic. However, it is also reasonable to consider using an effective AIDS vaccine in post-exposure immunization as immunotherapy against the disease (Salk. 1987).

It is widely believed that the HIV-1 envelope is the most promising candidate in the development of an AIDS vaccine (Francis and Petricciani. 1985; Vogt and Hirsh. 1986; Fauci. 1986). The HIV-1 envelope protein is initially synthesized as a 160,000 molecular weight glycoprotein (gp160). The gp160 precursor is then cleaved into a 120,000 molecular weight external glycoprotein (gp120) and a 41,000 molecular weight transmembrane glycoprotein (gp41). These envelope proteins are the major target antigens for antibodies in AIDS patients (Barin, et al. 1985). The native HIV-1 gp120 has been shown to be immunogenic and capable of inducing neutralizing antibodies in rodents, goats, rhesus monkeys and chimpanzees (Robey, et al. 1986).

Due to the very low levels of native HIV-1 envelope protein in infected cells and the risks associated with preparing an AIDS vaccine from HIV-1 infected cells, recombinant DNA methods have been employed to produce HIV-1 envelope antigens for use as AIDS vaccines. Recombinant DNA technology appears to provide the best option for the production of an AIDS subunit vaccine because of the ability to produce large quantities of safe and economical immunogens. The HIV-1 envelope has been expressed in genetically altered vaccinia virus recombinants (Chakrabarty et al. 1986; Hu et al. 1986; Kieny et al. 1986), bacterial cells (Putney et al. 1986), mammalian cells (Lasky, et al. 1986), and in insect cells. Synthetic peptides derived from amino acids sequences in an HIV-1 gp41 have also been considered as candidate AIDS vaccines (Kennedy, et al. 1986).

The use of a baculovirus-insect cell vector system to produce recombinant HIV-1 envelope proteins is disclosed in copending, coassigned U.S patent application Serial No. 920,197 filed October 16, 1986. The above-identified patent application and the publications referenced herein are herein incorporated and made part of this disclosure.

This system has been demonstrated to be of general utility in producing HIV-1 proteins and other proteins. As examples, the baculovirus Autographa californica nuclear polyhedrosis virus (AcNPV) has been used as a vector for the expression of the full length gp160 and various portions of the HIV-1 envelope gene in infected Spodoptera frugiperda (fall armyworm) cells (Sf9 cells). Also disclosed in the above-identified patent application is the truncated gp160 gene (recombinant number Ac3046), the protein produced from recombinant Ac3046, and a purification technique for the Ac3046 gene product that includes lentil lectin affinity chromatography followed by gel filtration chromatography. The gp160 protein purified in this manner and aggregated to form particles was found to be highly immunogenic in rodent and primate species.

The ideal AIDS vaccine, in addition to the requirements of being substantially biologically pure and non-pyrogenic, should provide life-long protection against infection with HIV-1 after a single or a few injections. This is usually the case with live attenuated vaccines. When killed bacteria or viruses, or materials isolated from them, such as toxoids or proteins, are used to make a vaccine, there often results a poor antibody response and only short term immunity. To overcome these deficiencies in a vaccine, an additional component, called an adjuvant, is added which has the property of being able to help stimulate the immune response. An adjuvant in common use in human vaccines (Bomford, 1985) are gels of salts of aluminum

(aluminum phosphate or aluminum hydroxide) and is usually referred to as an alum adjuvant.

### SUMMARY OF THE INVENTION

5

It has been discovered that recombinant HIV-1 gp160 protein, especially when adsorbed onto an adjuvant, such as alum, e.g. aluminum phosphate, is particularly useful as an AIDS vaccine. One aspect of this invention is an AcNPV expression vector having the coding sequence for a portion of the HIV-1 env gene which encompasses the amino acids 1-757 found in the recombinant clone number 3046 and the 10 production of that recombinant HIV-1 envelope protein in insect cells codes for by the amino acid sequences 1-757.

Other aspects of this invention comprise formation of recombinant envelope protein particles from the gene product from the baculovirus recombinant virus that produces the 3046 protein and adsorption of the 3046 particles to aggregates of aluminum phosphate.

15

### DETAILED DESCRIPTION OF THE INVENTION

The recombinant baculovirus *Autographa californica* nuclear polyhedrosis virus (AcNPV) which contains 20 the truncated HIV-1 gp160 gene coding for the HIV envelope protein 1-757 amino acids (recombinant Ac3046) is described in copending, coassigned U.S. application Serial No. 920,197, as well as the cloning step employed to construct the recombinant baculovirus containing genes or portions of genes from HIV-1. The following is a detailed description of the genetic engineering steps used to construct Ac3046 expression vector.

25 In the following description, the materials employed, including enzymes and immunological reagents, were obtained from commercial sources.

Details of the practices of this invention are set forth hereinbelow with reference to the accompanying drawings wherein:

Fig. 1 illustrates the cloning strategy used to isolate the HIV-1 envelope gene (env) from the E. coli 30 plasmid pNA2. The hatched regions are HIV-1 DNA sequences and the open regions are from the cloning vectors. The black region in p1774 is constructed from synthetic oligonucleotides and was introduced as a SmaI-KpnI fragment into the SmaI-KpnI sites of p1614. The sequence of this synthetic oligonucleotide is shown.

Fig. 2 illustrates the strategy used to construct the recombinant vector (p3046) used to construct the 35 baculovirus expression vector Ac3046. The plasmid pMGS3 contains sequences (cross-hatched areas) from the baculovirus AcNPV on either side of a cloning site at position 4.00. This site has the unique restriction endonuclease sites for SmaI, KpnI, and BglII. In the 5' direction from the 4.00 position is the AcNPV polyhedrin promoter and in the 3' direction is the sequence 5'-TAATTAATTAA-3' which has a translational codon in all three reading frames. The plasmid p1774 and the sequence of the synthetic oligonucleotide 40 region is as described in Fig. 1. The plasmid p3046 contains all of pMGS3 except for the sequence between the SmaI and BglII sites and the HIV-1 envelope gene in p1774.

Fig. 3 shows the nucleotide sequences of the DNA flanking the Ac3046 gp160 coding sequences. The 3046 env DNA sequence between +1 and +2264 is shown in Fig. 4.

Figs. 4a-4k show the actual DNA sequence of the HIV-1 env gene segment along with the synthetic 45 oligonucleotide sequences at the 5' end of the env gene in Ac3046 between +1 to +2264. The location of restriction endonuclease sites are listed above the DNA sequence and the predicted amino acid sequence is listed below the DNA sequence. The bases are numbered on the right and on the left.

Fig. 5a-5d compare the DNA sequences of the env gene from Ac3046 with a published env gene 50 sequence from LAV-1. The LAV-1 sequence is on the top and Ac3046 is on the bottom. A line ( ) below the LAV-1 sequence indicates that the sequence in Ac3046 is the same in this position. The DNA sequence numbering used is that described by Wain-Hobson, et al 1985 for LAV-1.

Fig. 6 shows the ELISA end point dilution titers of human HIV-1 antibody positive sera (top graph) and rhesus monkey sera (bottom graph) from animals immunized with gp160 (IJ55, KL55) or gp120 (AB55, 55 CD55, GH55). The ELISA titers were measured against highly purified gp120 and gp160 proteins. The specifically bound antibody was measured with a goat anti-human IgG HRP conjugate. The highest dilution of serum that gives a positive response in the test is the titer.

The following are examples illustrative of the practices of this invention:

EXAMPLE 1

Construction of the baculovirus recombinant Ac3046 bearing the HIV-1 coding sequence for amino acids 1-  
 5 757

Cloning and expression of foreign protein coding sequences in a baculovirus vector require that the coding sequence be aligned with the polyhedrin promoter and upstream sequences and on the other side with truncated polyhedrin coding sequences such that homologous recombination with the baculovirus genome results in transfer of the foreign coding sequence aligned with the polyhedrin promoter and an inactive polyhedrin gene.

Accordingly, a variety of insertion vectors were designed for use in AIDS env gene constructions. The insertion vector MGS-3, described below, was designed to supply the ATG translational initiating codon. Insertion of foreign sequences into this vector must be engineered such that the translational frame established by the initiating codon is maintained correctly through the foreign sequences.

The insertion vector MGS-3 was constructed from an EcoRI-I restriction fragment clone of DNA isolated from a plaque purified AcMNPV isolate (WT-1), and was designed to consist of the following structural features: 4000 bp of sequence upstream from the ATG initiating codon of the polyhedrin gene; a polylinker introduced by site directed mutagenesis, which consists of an ATG initiating codon at the exact position of the corresponding polyhedrin codon, and restriction sites SmaI, KpnI, BglII and a universal stop codon segment; 1700 bp of sequence extending from the KpnI restriction site (which is internal to the polyhedrin gene) through to the terminal EcoRI restriction site of the EcoRI-I clone.

EXAMPLE 2

Construction of baculovirus recombinants bearing LAV env coding sequences

A recombinant plasmid designated NA-2 consisted of a 21.8 kb segment of an entire AIDS provirus inserted into pUC18. This clone was reportedly infectious since it could produce virus following transfection of certain human cells (Adachi et al. 1986). The complete envelope gene sequences contained in NA-2 were derived from the LAV strain of HIV (Barre-Sinoussi et al. 1983).

The HIV-1 envelope gene was isolated and engineered as described below, see also Fig. 1. The envelope gene was initially isolated from NA-2 as a 3846 bp EcoI/SacI restriction fragment and cloned into the EcoRI/SacI restriction site pUC19. The resultant plasmid was designated as p708. The envelope gene was subsequently reisolated as a 2800 bp KpnI restriction fragment and cloned into the KpnI restriction site of pUC18. The resulting clone was designated p1614. This KpnI restriction fragment contained a slightly truncated piece of the envelope gene such that 121 bp of the N-terminal corresponding sequence was missing. This missing part in the gene, which included the signal peptide sequences, was replaced by insertion of a double-stranded synthetic oligomer which was designed from the LAV amino acid sequence using preferred polyhedrin gene codon usage. To facilitate further manipulation, a new SmaI restriction sequence was concomitantly introduced in place of the ATG initiating codon. The ATG initiation codon will be supplied by the insertion vector. The resultant plasmid was designated a p1774.

Referring now to Fig. 2, restriction fragments from p1774 containing coding sequences of various domains of the AIDS envelope were cloned into the MGS vectors such that the ATG initiating codon of the insertion vector was in-frame with the codons of the envelope gene. Construct p3046 consisted of the SmaI/BamHI restriction fragment isolated from p1447 inserted into the SmaI/BglII site of the plasmid vector pMGS-3. This clone contains sequences coding for amino acids 1 through 757 of gp160 and uses a termination codon supplied by the MGS-3 vector.

EXAMPLE 3

55

Preparation and Selection of Recombinant Baculovirus

The HIV env gene recombination plasmid p3046 was calcium phosphate precipitated with AcMNPV DNA

(WT-1) and added to uninfected Spodoptera frugiperda cells. The chimeric gene was then inserted into the AcMNPV genome by homologous recombination. Recombinant viruses identified were by an occlusion negative plaque morphology. Such plaques exhibit an identifiable cytopathic effect but no nuclear occlusions. Two additional successive plaque purifications were carried out to obtain pure recombinant virus.

5 Recombinant viral-DNA was analyzed for site-specific insertion of the HIV env sequences by comparing their restrictions and hybridization characteristics to wild-type viral DNA.

#### EXAMPLE 4

10

##### Expression of HIV env from recombinant baculoviruses in infected insert cells

Expression of HIV-env sequences from the recombinant viruses in insect cells should result in the synthesis of primary translational product in the form of a pre-pro-protein containing all the amino acids coded for from the ATG initiating codon of the expression vector downstream from the polyhedrin promoter. This primary product will consist of amino acids translated from the codons supplied by the recombination vector. The primary translation product of Ac3046 should read Met-Pro-Gly-Arg-Val at the terminus where Arg (position 4) is the Arg at position 2 in the original LAV clone. The Met-Pro-Gly codons are supplied as a result of the cloning strategy.

#### EXAMPLE 5

25

##### Nucleotide sequence of the gp160 insert and flanking DNA

The nucleotide sequence of the gp160 insert and flanking DNA was determined from restriction fragments isolated from viral expression vector Ac3046 DNA. The sequencing strategy involved the following steps. The 3.9 kbp EcoRV-BamHI fragment was purified by restriction digestion of Ac3046 viral DNA. The Ac3046 viral DNA had been prepared from extracellular virus present in the media of cells being used for a production lot of vaccine.

As shown in Fig. 2, the 3.9 kbp EcoRV-BamHI fragment consists of the entire gp160 gene and 100 bp of upstream and about 1000 bp of downstream flanking DNA. Of this, the nucleotide sequence of the entire gp160 gene was determined, including 100 bp of upstream and 100 bp of downstream flanking DNA.

Briefly, the results of the sequencing revealed a chimeric construct as predicted from the cloning strategy. The sequence of the gp160 was essentially as reported by Wain-Hobson et al. (1985). The sequence of 2253 bases between the presumed translation initiation and termination codons predicts 751 amino acid codons and 28 potential N-linked glycosylation sites. The estimated molecular weight of gp160, including the sugar residues, is approximately 145,000.

Sequence analysis of 200 bases of flanking DNA indicated correct insertion as indicated in Figs. 3, 4 and 5.

#### EXAMPLE 6

##### Amino Acid Sequence of gp160

Using standard automated Edman degradation and HPLC procedures, the N-terminal sequence of the first 15 residues of gp160 was determined to be identical to that predicted from the DNA sequence. The N-terminal methionine is not present on the gp160 protein. This is consistent with the observation that AcNPV polyhedrin protein is also produced without an N-terminal methionine. A summary of the actual 5' gp160 DNA and N-terminal protein sequences, as has been determined by analysis of the AcNPV 3046 DNA and purified gp160, is as follows:

##### LAV env gene in the AcNPV 3046 expression vector

Residue # 2 3 4 5 6 7 8 9 10 11 12 13 14  
 - Pro Gly Arg Val Lys Glu Lys Tyr Gln His Leu Trp Arg  
 Trp Gly  
 5 ATG CCC GGG CGT GTG AAG GAG AAG TAC CAA CAC CTG TGG  
 CGT TGG GGC

10

Lav env gene in the original LAV-1 clone

Residue # 1 2 3 4 5 6 7 8 9 10 11 12 13 14  
 15 Met Arg Val Lys Glu Lys Tyr Gln His Leu Trp Arg Trp  
 Gly  
 ATG AGA GTG AAG GAG AAG TAT CAG CAC TTG TGG AGA TGG  
 20 GGG

25 EXAMPLE 7Purifications of recombinant gp160

30 One aspect of the prevent invention is the procedure used to extract and purify the recombinant HIV-1 envelope protein coded for in the Ac3046 expression vector. The recombinant HIV-1 envelope protein gp160 is produced in S. frugiperda cells during the 4-5 days after infection with Ac3046. Purification of the gp160 protein involves the steps:

1. Washing the Cells
2. Cell Lysis
3. Lentil Lectin Affinity Chromatography
4. Gel Filtration Chromatography
5. Dialysis

40 The following steps describe the purification of the recombinant gp 160 from Ac3046 infected cells obtained from 2 x 10<sup>9</sup> infected cells:

Washing the cells.

45 Infected cells are washed in a buffer containing 50 mM Tris buffer, pH 7.5, 1 mM EDTA and 1% Triton x-100. The cells are resuspended in this buffer, homogenized using standard methods, and centrifuged at 5000 rpm for 20 min. This process is repeated 3 times.

50 Cell Lysis.

After the washed cells are lysed by sonication in 50 mM Tris buffer, pH 8.0-8.5, 4% deoxycholate, 1% B-mercaptoethanol. Sonication is done using standard methods. After sonication, only remnants of the nuclear membrane are intact and these are removed by the centrifugation of 5000 rpm for 30 min. The supernatant containing the extracted gp160 contains no intact cells as determined by light microscopy observations.

Gel filtration.

Gel filtration is done in a Pharmacia 5.0 x 50 cm glass column packed with a Sephadryl resin (Pharmacia). The total bed volume is about 1750 ml. To depyrogenate and sanitize the column and tubing connections, at least 6 liters of 0.1 N NaOH is run through the column over a period of 24 hours. The effluent from the column is connected to a UV flow cell and monitor and chart recorder (Pharmacia) then equilibrated with 4 liters of Gel Filtration Buffer. The crude gp 160 is loaded onto the column then developed with Gel Filtration Buffer. The column separates the crude mixture into three major UV absorbing fractions. The first peak comes off between about 500 and 700 mls, the second between 700 and 1400 mls and the third between 1400 and 1900 mls buffer. This same profile is observed on small analytical columns from which it has been determined that the first peak is material that has a molecular weight of  $\geq 2,000,000$ . This peak is translucent due to a concentration of high molecular weight lipids and lipid complexes. This peak also contains from 10% to 20% of the gp 160 extracted from the infected cells. Apparently this fraction of gp160 is complexed to itself or other cell components to form high molecular weight aggregates. The second broad peak contains the majority of the gp160 and proteins with molecular weights of between about 18,000 and 200,000. The third peak contains little protein and the majority of the UV absorption is due to the B-mercaptoethanol in the sample. When the second peak is first detected from the tracing of the UV absorbance, the effluent from the column is directly onto the lentil lectin column. Once the second peak has come off the column, the effluent is disconnected from the lentil lectin column and directed to waste.

20

Lentil lectin.

The lentil lectin affinity gel media (Lentil Lectin-Sepharose 4B) was purchased in bulk from Pharmacia. The lentil lectin was isolated by affinity chromatography on Sephadex to greater than 98% purity then immobilized by coupling to Sepharose 4B using cyanogen bromide. The matrix contains about 2 mg ligand per ml of gel. The lentil lectin column is a 5.0 x 30 cm glass column (Pharmacia) containing 125 ml lentil lectin-Sepharose 4B gel. The affinity matrix is reused after being thoroughly washed and regenerated by a procedure recommended by the supplier. When not in use, the gel is stored in the column in a solution of 0.9% NaCl, 1 mM MnCl<sub>2</sub>, 1 mM CaCl<sub>2</sub>, and 0.01% thimerasol. The column is washed and equilibrated with 250 ml lentil lectin buffer described above before each use. The crude gp160 is applied to the column directly as it is eluting from the gel filtration column as described above. Once the crude gp160 is bound to the column, it is washed with 800 ml lentil lectin buffer containing 0.1% deoxycholate. Under these conditions all of the gp160 binds to the column. lentil lectin buffer plus 0.3M alpha-methyl mannoside is used to elute the bound glycoproteins which is monitored through a UV monitor at a wavelength of 280 nm.

Purification summary.

40      Summary of the purification of gp 160 from 1 liter of infected cells:

Purification Step	Total Protein <sup>1</sup> mg	gp160 Protein mg	% gp160 Total	Contaminants Removed
Cell pellet	1-2000	20	1-2	Culture medium
1,2,3rd Wash	250	15	6	Serum Albumin, most Nucleic Acids, and Soluble Cell Proteins
Gel Filtration	120	12	10	Lipids, Nucleic Acids, and high mol wt aggregates
Lentil Lectin	14	10	70	Nonglycosylated proteins
Dialysis	13	9	70	Sugar, deoxycholate, excess Tris buffer

<sup>1</sup> The total protein content was estimated from the absorbance at 280 nm

55

EXAMPLE 8

Assembly of gp160 Particles.

As one aspect of the present invention, it has been discovered that the gp160 antigen can be assembled into particles of  $\geq 2,000,000$  molecular weight during purification. The gp160 protein is extracted from the cell as a mixture of 80-90% monomeric (160,000 molecular weight) and 10-20% polymeric (particle form). The gel filtration step removes the aggregated forms of gp160. Attempts to purify the gp160 from this fraction (first peak off the gel filtration column) suggest that it is complexed with other cell proteins, possibly even with membrane fragments. However, the gp160 antigen in the second peak off the gel filtration column has a molecular weight of about 160,000-300,000 and is, therefore, in predominantly monomeric or dimeric form. The formation of aggregates of polymers of gp 160 occurs during the development of the lentil lectin column. It has been determined that the antigen forms aggregates whether it is eluted from the lectin column in 0.5% doxycholate, which is about the 0.2% critical micelle concentration (CMI) for doxycholate, or whether the gp160 is eluted from the column in 0.1% deoxycholate. The size of the aggregates are measured on a high resolution FPLC Superose 12 column (Pharmacia). Samples from representative lots of purified gp160 have a size that is predominantly equal to or greater than the 2,000,000 molecular weight of a blue dextran size standard. It is likely that, as non-glycosylated proteins are removed from association with the gp160 antigen during the binding and washing to the lentil lectin column, the hydrophobic portions of gp160 begin to form intermolecular associations. The deoxycholate is probably not bound to the gp160 as the concentration can be kept above the CMI and the antigen will still form complexes. The assembly of this antigen into aggregates appears to be an intrinsic property of this protein once it is purified. It is possible that the very hydrophobic N-terminal sequence that is present on the gp160 protein contributes to the natural ability of this protein to form particles. After purification, the gp160 complexes can be sterile filtered through an 0.2 micron cellulose acetate filter without significant loss of protein.

25

Analysis of Particle Formation.

An analysis of purified gp160 by electron microscopy demonstrates that there were protein-like, spherical particles of 30-100 nM. As an additional test for the presence of particles, purified gp160 was analyzed by gel filtration. About 100  $\mu\text{g}$  gp160 was applied to a Superose 12, FPLC gel filtration HR 10/30 column (Pharmacia, Inc.). This column was first calibrated with protein molecular weight standards. The protein profile from this column is highly reproducible with respect to the elution volume being inversely proportional to the molecular weight of protein standards. The column separates the monomeric gp160 from the polymeric forms and excludes globular proteins of  $\geq 2 \times 10^6$  molecular weight. When developed on this column, essentially all of the purified gp160 elutes in the void volume and is, therefore,  $\geq 2 \times 10^6$  molecular weight in size.

40 EXAMPLE 9Adsorption of gp160 to Alum.

45 The effectiveness of insoluble aluminum compounds as immunologic adjuvants depends on the completeness of adsorption of the antigens on the solid phase. As a part of the present disclosure it was discovered that alum could be made that would efficiently adsorb the gp160 but at a pH that would not reduce the potency of the gp160-alum complex as an immunogen. The factors controlled during the formation of alum (aluminum phosphate gel) are:

- 50 1. The optimal pH for adsorption of antigens to alum is about 5.0. However, it was discovered that the gp160 lost immunogenicity at a pH of 6.5 in comparison to a pH of 7.5 so the Alum is made at a pH of 7.1  $\pm 0.1$ . It was discovered that essentially 100% of the gp160 will still adsorb to the alum at this pH.  
 2. The ionic strength from the NaCl present is relatively low and is less than 0.15 M.  
 3. There is a molar excess of aluminum chloride relative to sodium phosphate to assure that there is an absence of free phosphate ions in the supernatant.  
 55 4. The gp160 antigen is added to freshly formed alum to stop crystal growth and minimize the size of the particles.

The procedure to make 100 ml alum and adsorb purified gp160 to the alum so that the final concentration of antigen is 40 µg/ml is outlined hereinbelow:

5 Preparation of Reagents (200 mL total formulated lot).

10 1. Prepare the following solutions in 100 mL sterile, pyrogen-free bottles or beakers. Mix the salts for Solution 1 and Solution 2 and the sodium hydroxide and filter through 0.2µ cellulose acetate filters into 100 mL sterile, pyrogen-free bottles.

15

Solution 1	AlCl3.6H2O NaHAc.3H2O	0.895 grams 0.136 grams
dissolve in 40 mL WFI water, 0.2µ filter		
Solution 2	Na3PO4.12H2O	1.234 grams
dissolve in 40 mL WFI water, 0.2µ filter		
Solution 3	NaOH	2.0 grams
dissolve in 100 mL WFI water, 0.2µ filter		
Solution 4	Tris	1.25 grams
dissolve in 100 mL WFI water, add 1 mL to 90 mL WFI, adjust pH to 7.5 with 0.5N HCl, and bring to 100 mL with WFI		

20 2. Autoclave solutions for 30 min; slow exhaust. Cool to room temperature.

25

Formation of Alum

30 1. Add Solution 1 (aluminum chloride-sodium acetate to the formulation vessel using 25 mL sterile, disposable pipets. Note the volume of Solution 1 and begin stirring the solution.

35

2. Add Solution 2 (sodium phosphate) to the vessel using 25 mL sterile, disposable pipets and continue stirring as the precipitate forms and note the volume of Solution 2.

3. Add 3mL Solution 3 (sodium hydroxide) and continue stirring for 5 min. Take an 0.5 ml sample and measure the pH. If the pH is less than 7.0, add an additional 0.5 ml sodium hydroxide, stir for another 5 minutes and measure the pH again. Continue until the pH is between 7.0 and 7.2.

40

4. Determine the total volume added to the formulation vessel (Solution 1 + Solution 2 + Solution 3), then add sterile WFI to bring the volume to 100 mL.

5. Immediately add 8,000 µg purified gp160 in 100mL of 1mM Tris pH 7.5 directly into the formulation vessel.

45

6. Continue stirring for a minimum of 20 minutes, then dispense the formulated vaccine into a sterile vial.

EXAMPLE 10

50

Immunogenicity of Alum adsorbed gp160.

55

An accepted method to determine the immunogenicity of an antigen preparation (vaccine) is to measure the specific antibody response in groups of mice which have been given a single dose of antigen. At the end of 4 weeks the mice are bled and the serum antibody levels to a specified antigen (usually the antigen used to immunize the animal with) measured by a standard antibody test such as an ELISA (enzyme linked

immunoabsorbant assay).

The immunogenicity in mice of purified gp160 with no adjuvant at pH 6.0 and pH 7.5 adsorbed as described in Example 9 with alum, or mixed with Freunds Complete Adjuvant are summarized hereinbelow:

5	Group		gp160		Seroconversion	
	gp160	Adjuvant	Lot#	Mean ELISA OD <sup>1</sup>	%	(P/N) <sup>2</sup>
10	1 $\mu$ g	None, pH 7.5	8702	0.140	57%	4/6
		None, pH 6.0	8702	0.110	26%	2/7
		Alum	8702	1.000	90%	9/10
		Alum	8705	2.285	100%	6/6
		Freund's	8604	1.108	83%	5/6
		Freund's	8702	1.396	100%	7/7
15	0.1 $\mu$ g	Freund's	8604	0.434	67%	4/6
			8705	1.003	67%	4/6

<sup>1</sup> The mice were bled 28 days post immunization and the sera tested at 1:10 dilution in an ELISA assay against gel-purified gp160. Similar results were obtained using a commercial ELISA (Genetic Systems Inc.; EIA<sup>TM</sup> ELISA) assay against the native HIV-1 proteins at a serum dilution of 1:400.

<sup>2</sup> The number of seroconverted mice (P) to the total number tested (N).

25 Mice immunized with single 1.0  $\mu$ g dose of gp160 antigen without any added adjuvant will elicit an antibody response against gp160 (see table above). However, a much stronger antibody response is seen in groups of mice immunized with 1.0  $\mu$ g gp160 adsorbed to alum adjuvant. A single dose of less than 0.1  $\mu$ g of gp160 mixed with complete Freund's or formulated with alum will seroconvert  $\geq$  50% of the immunized mice. Although less so, the gp160 antigen was immunogenic in mice as an unformulated antigen at a pH 7.5 and a pH 6.0, but there was a loss of immunogenicity at the lower pH.

30 The ability of a candidate vaccine to elicit an immune response is a very important biological property. To confirm that the alum formulated gp160 vaccine was immunogenic in animals and to confirm that the alum adjuvant increased this immunogenicity, the following experiment was performed. On day 0, mice (groups of 10) were injected with a single dose (0.5  $\mu$ g or 5.0  $\mu$ g) of gp160 alone, gp160 adsorbed to alum or gp160 in complete Freund's adjuvant (CFA). On day 28 the mice were bled and the sera examined by ELISA (1:10 dilution) for the presence of antibodies to gp160.

35 Results from the sera drawn on day 28 are summarized in the table below. In all groups, greater than 50% of the mice showed seroconversion. At all the doses the number of seroconversions and the average serum absorbance readings (OD<sub>450 nm</sub> at a 1:10 dilution in the ELISA assay) were higher with gp160 adsorbed to alum than those obtained in mice immunized with gp160 alone. These results demonstrate that the alum adjuvant significantly increased the immunogenicity of the gp160 antigen.

45

50

55

28 Days Post-Injection

	0.5 µg Dose		1.0 µg Dose		5.0 µg Dose		
	P/N <sup>1</sup>	Mean OD <sup>2</sup>	P/N	Mean OD	P/N	Mean OD	
10	gp160	9/10	.407	7/10	.699	7/10	.430
15	gp160 (alum)	9/10	.547	8/10	.797	10/10	1.347
15	gp160 (CFA)	10/10	1.130	10/10	1.967	10/10	1.317

20           <sup>1</sup> The number of mice that seroconverted (P) compared to total number tested (n) at 28 days after being immunized with 0.5 µg, 1 µg or 5 µg of VaxSyn™ HIV-1.

25           <sup>2</sup> The mean absorbance (OD 450) of the mice that seroconverted as measured by the sponsor's ELISA assay against gp160 at a 1:10 dilution of serum

35           **EXAMPLE 12**

Neutralization Data

40           HIV-1 neutralization assays are an accepted method to determine whether an antibody preparation will inhibit the HIV-1 virus from infecting susceptible human cultured lymphocyte cells. Antisera from animals immunized with gp160 were tested in an HIV-1 neutralization assay and the results summarized in the table below.

45

50

55

Animal	Identification	Immunogen/Adjuvant	$\mu\text{g}^1$	Neutralizing Titer <sup>2</sup>
Rhesus	G55	gp120/Alum	16/8/8	1:80-1:160
Rhesus	H55	gp120/Alum	16/8/8	1:80-1:160
Rhesus	L55	gp160/Alum	16/8/8	$\geq 1:80$
Mice	Pool 3	gp120/Freunds	.25/.25.25	1:40-1:80
Mice	Pool 8	gp160/Freunds	.1/.1/.1	1:40-1:80
G. Pig	purified IgG	gp160/Freunds	10/10/10	1:320

<sup>10</sup> <sup>1</sup> Micrograms of gp160 or gp120 administered during the first/second/third immunization.

<sup>2</sup> <sup>2</sup> The highest dilution of antisera that will inhibit the infection by 50% relative to HIV-1 infected cells that were exposed to serum from non-immunized animals

<sup>15</sup> Guinea pigs, rabbits and rhesus monkeys have also been immunized with gp160 (using alum or Freunds as an adjuvant). In general, the immunization of these animals has produced a good antibody response against the HIV-1 envelope proteins.

<sup>20</sup> **EXAMPLE 13**

Immunogenicity in Chimpanzees

<sup>25</sup> Genetically, the chimpanzee is man's closest relative and is currently the only animal model for infection of HIV-1. In a safety/immunogenicity trial in three chimpanzees, two chimpanzees were immunized with 40  $\mu\text{g}$  or 80  $\mu\text{g}$  gp160 in alum formulated vaccine. Each received a booster immunization at 4 weeks with 40  $\mu\text{g}$  and 80  $\mu\text{g}$  gp 160, respectively. A control animal was vaccinated at the same time with 1 ml saline solution. Weekly serum samples were analyzed from each of the three chimpanzees for antibodies to gp160 and to HIV-1 viral antigens using three immunological assays, an ELISA assay against purified gp160 developed by MicroGeneSys, Inc., Western Blot analysis, and a commercial HIV-1 ELISA assay. The results of these analyses are described below:

<sup>35</sup> ELISA (MGSearch HIV 160).

The ELISA assay, MGSearch HIV 160, MGSearch being a trademark of MicroGeneSys, Inc. of West Haven, Connecticut, U.S.A., is an immunoassay against gp160 and is described in copending coassigned U.S. patent Application Serial No. 920,197. Serum samples taken before immunization and for the 11 weeks following the primary immunization were diluted from 1:10 to 1:100,000 and then incubated with nitrocellulose strips containing 100 ng purified gp160 in a spot. The end point dilution titer is the highest dilution in which the test was positive for anti-gp160 antibody as detected with a goat anti-human IgG-alkali phosphatase conjugate. The serum samples from the control animal and from the preimmune sera of the immunized animal were negative. The chimp which received a 80  $\mu\text{g}$  dose was positive at a 1:100 dilution by week 2 and the chimp which received a 40  $\mu\text{g}$  dose was positive at 1:10 dilution by week 4. The antibody titers to gp160 continued to increase until week 5, at which time the end point dilution titers were approximately 1:100,000 and 1:2,000, respectively. The antibody titer in both animals dropped just slightly during weeks 6-11. This type of response is similar both quantitatively and qualitatively to antibody responses commonly observed in chimps that have been vaccinated with a human Hepatitis B Virus Vaccine.

Commercial ELISA Test

<sup>55</sup> It was clear from the MGSearch HIV 160 ELISA and Western blot analyses of sera from the VaxSyn immunized chimpanzees, VaxSyn being a trademark of MicroGeneSys, Inc. for AIDS vaccine, that they had seroconverted and have antibodies against the recombinant gp160. To determine if they were also making

anti-HIV antibody which recognized the native viral envelope proteins, the preimmune sera and sera from weeks 1 through 11 were tested in a licensed, commercial ELISA test kit LAV EIA test kit, LAV EIA being a trademark of Genetic System Corporation, Seattle, Washington, U.S.A. The animal immunized with 80 µg gp160 was positive at a 1:100 dilution by week 2 and continued to show an increase in antibody level through week 6. The animal immunized with 40 µg was positive at a 1:100 dilution by week 6.

#### EXAMPLE 14

10

##### Distribution of Antibodies Between gp120 and gp41

It is important to determine whether the antibody responses against gp160 in a vaccinated animal is directed against gp41 or gp120 or both. A variety of immunological methods, including radioimmuno-precipitation (RIP), immunofluorescence (IF), Western blot analysis, and quantitative ELISA against three different recombinant envelope antigens were employed to detect and measure for the distribution of antibodies against various regions of the HIV-1 envelope proteins. Fig 6 summarizes the immunoreactivity of three different recombinant antigens (1) gp120-γ (truncated recombinant HIV-1 gp120 with about 40 amino acids missing from the C-terminus of the molecule), (2) gp120 (full length recombinant HIV-1 gp120; 20 and (3) gp160. Human sera from 50 HIV-1 antibody positive individuals and 3 pooled human sera were highly reactive with gp160, moderately reactive with gp120 and little or no antibody reacted with truncated gp120. It is likely that the truncated gp120, which represents more than 90% of the HIV-1 external glycoprotein, contains protective determinants. The observation that human AIDS positive sera have few antibodies to this region of the envelope is consistent with the fact that the immune response to viral 25 infection is not fully protective and that human positive sera usually exhibit a low-level of neutralizing activity *in vitro*. In contrast, rhesus monkeys immunized with either the gp160 immunogen or with the truncated gp120 have antibodies that react strongly with the truncated gp 120 portion of the HIV-1 envelope. This difference in distribution of antibody recognition sites along the viral envelope and the higher titers observed in the monkeys may account for the fact that the monkey sera had high neutralizing titers. A quantitative 30 assessment of the immunoreactivity of these three recombinant envelope antigens with human and immune rhesus sera is presented in Fig. 7. All the monkey sera tested had high titer antibody against the truncated gp120 antigen (gp120-γ), including those from animals immunized with gp160.

The following is a listing of and a more complete identification of the reference cited hereinabove:

35

##### References

- Adachi, A. H.E. Gendelman, S. Koenig, T. Folks, R. Willey, A. Rabson, and M. Martin. 1986. Production of acquired immunodeficiency syndrome-associated retrovirus in human and nonhuman cells transfected with an infectious molecular clone *J. Virol.* 59:284-291.
- Barin F, MF McLane, JS Allan, et al. Virus envelope protein of HTLV-III represents major target antigen for antibodies in AIDS patient *Science*. 228:1094-1096.
- Barré-Sinoussi F, JC Chermann, F Rey, et al. 1983. Isolation of T-Lymphotropic Retrovirus from a Patient at Risk for Acquired Immune Deficiency Syndrome. *Science*. 220:868-871.
- Bomford R. 1985. Adjuvants. *in Animal Cell Biotech.* Vol. 2:235-250. Academic Press Inc. (London) Ltd.
- Chakrabarti, S, M Robert-Guroff, F Wong-Stall, RC Gallo, B Moss. 1986. Expression of the HTLV-III envelope gene by a recombinant vaccinia virus. *Nature*. 320:535-537.
- Cochran MA, BL Ericson, JD Knell, GE Smith. 1987. Use of baculovirus recombinants as a general method for the production of subunit vaccines. *in Vaccines 87*. Cold Springs Harbor Press, in press.
- Fauci, Anthony S. 1986. Current issues in developing a strategy for dealing with the acquired immunodeficiency syndrome. *Proc. Natl. Acad. Sci. USA*. 83:9278-9238.
- Feorino, PM, VS Kalyanaraman, HW Havercos, et al. 1984. Lymphadenopathy associated virus infection of a blood donor-recipient pair with acquired immunodeficiency syndrome. *Science*. 225:69-72.
- Francis DP, JC Petricciani. 1985. The prospects for and pathways toward a vaccine for AIDS. *New Eng. Journal of Med.* 1586-1590.
- Kennedy RC, RD Henkel, D Pauleti, JS Allan, TH Lee, M Essex, GR Dreesman. 1986. Antiserum peptide recognizes the HTLV-III envelope glycoprotein *Science*. 231:1556-1559.

- Kieny MP, G Rautmann, D Schmitt, K Dott, S Wain-Hobson, M Alikzon, M Girard, S Chamaret, A Laurent, L Montagnier, and J-P Lecocq. 1986. AIDS virus *env* protein expressed from a recombinant vaccinia virus. **Biotechnology**. 4:790-795.
- Hu, S-L, SG Kosowski, JM Dalrymple. 1986. Expression of AIDS virus envelope gene in recombinant 5 vaccinia viruses. **Nature**. 320:537-540.
- Lasky, LA JE Groopman, CW Fennie, PM Benz, DJ Capon, DJ Dowabenko, GR Nakamura, WM Nunes, ME Renz PW Berman. 1986. **Science**. 233:209-212.
- Levy, JA, AD Hoffman, SM Kramer, JA Landis, JM Shimabukurs, and LS Oshiro. 1984. Isolation of lymphcytopathic retroviruses from San Francisco patients with AIDS. **Science**. 225:840-842.
- Popovic, M, M Sarangadharan, E Read, and R Gallo. 1984. Detection, Isolation and Continuous 10 Production of Cytopathic Retroviruses (HTLV-III) from patients with AIDS and at risk for AIDS. **Science**. 224:497-500.
- Robey WG, LO Arthur, TJ Matthews *et al.* 1986. Prospect for prevention of human immunodeficiency virus infection: purified 120-kDa envelope glycoprotein induces neutralizing antibody. **Proc. Natl. Acad. 15 USA**. 83:7023-7027.
- Salk, J. 1987. Prospects for the control of AIDS through post-exposure immunization. **Nature**. 327:473-476.
- Putney, SD., Matthews, T.J., Robey, W.G., Lynn, D.L., Robert-Guroff, M., Muller, W.T., Langlois, A.J., Ghrayeb, J., Petteway, S.R., Jr., Weinhold, K.J., Fischinger, P.J., Wong-Staal, F., Gallo, R.C., and 20 Bolognesi, D.P. 1986. HTLV-III/LAV Neutralizing antibodies to an *E. coli*-produced fragment of the virus envelope. **Science**. 234:1392-1395.
- Vogt, M. and M.S. Hirsch. 1986. Prospects for the prevention and therapy of infections with the human immunodeficiency virus. **Reviews of Infectious Disease**. 8:991-1000.
- Wain-Hobson, S., Sonigo, P., Danos, O., Cole, S., and Alizon, M. Nucleotide Sequence of the AIDS 25 virus, LAV. 1985. **Cell** 40:9-17.

### Claims

1. Undenatured or native HIV-1 envelope protein.
2. Recombinant undenatured or native HIV-1 envelope protein of Claim 1.
3. Substantially biologically pure HIV-1 envelope protein in accordance with Claim 1.
4. Substantially biologically pure HIV-1 envelope protein in accordance with Claim 1 of a purity and 35 quality useful as a vaccine for humans.
5. Substantially biologically pure gp160 protein in accordance with Claim 1.
6. Recombinant substantially pure HIV-1 envelope protein in accordance with Claim 1 of a purity and quality useful as a vaccine for humans.
7. A substantially non-pyrogenic vaccine composition containing a protein in accordance with Claim 1 40 together with an adjuvant.
8. A composition in accordance with Claim 7 wherein said adjuvant is particle-form aluminum phosphate.
9. A composition in accordance with Claim 7 wherein said adjuvant is particle-form aluminum hydroxide.
10. A composition for the treatment of acquired immunodeficiency syndrome (AIDS) comprising a 45 minor amount of a protein in accordance with Claim 1 together with a carrier.
11. gp160 HIV-1 envelope protein in accordance with Claim 1.
12. Recombinant gp160 HIV-1 envelope protein in accordance with Claim 1.
13. A recombinant DNA molecule for expressing HIV-1 envelope protein in accordance with Claim 1.
14. A DNA molecule of Figs. 4a-4k herein for expressing HIV-1 envelope protein in accordance with 50 Claim 1.
15. A recombinant DNA molecule containing the DNA sequence for expressing the HIV-1 envelope protein of Claim 1.
16. A recombinant DNA molecule in accordance with Claim 15 containing a DNA sequence for expressing gp160 HIV-1 envelope protein.
17. A DNA molecule in accordance with AC3046 herein.
18. Substantially biologically pure undenatured or native HIV-1 envelope protein in accordance with 55 Claim 1 recovered after expression in a host insect cell.

19. Substantially biologically pure undenatured or native gp160 HIV-1 envelope protein in accordance with Claim 18.
20. Recombinant substantially biologically pure undenatured or native HIV-1 envelope protein expressed by Ac3046 herein in accordance with Claim 18.
- 5 21. The recombinant DNA molecule Ac3046 herein in accordance with Claim 17 for the expression of undenatured or native gp160 envelope protein of HIV-1 in an insect cell.
22. A method of treating a human patient suffering from AIDS which comprises administering to the patient a therapeutically effective amount of a protein in accordance with Claim 1.
- 10 23. A method for preventing AIDS in a human which comprises administering to said human an effective prophylactic amount of a protein in accordance with Claim 4.
24. A method of recovering and purifying recombinant HIV-1 envelope protein material expressed in a cell which comprises after expression of said protein material recovering and washing the cells, lysing the cells, subjecting the resulting lysate to a lectin affinity chromatography followed by gel filtration chromatography and dialysis and recovering the resulting purified HIV-1 envelope protein material.
- 15 25. A method in accordance with Claim 24 wherein said cells containing the to-be-recovered HIV-1 envelope protein material therein are subjected to washing in a buffer solution containing 50 nM Tris buffer at a pH 7.5 and 1nM EDTA and 1% Triton X-100.
26. A method in accordance with Claim 24 wherein cells lysis is carried out by sonication in a 50 nM Tris buffer at a pH in the range about 8.0-8.5 and 4% deoxycholate and 1% beta mercaptoethanol.
- 20 27. A method in accordance with Claim 24 wherein said gel filtration is carried out to a column containing Sephadex resin.
28. A method in accordance with Claim 24 wherein said lectin affinity chromatography is carried out employing lentil lectin.
29. A method in accordance with Claim 24 wherein the resulting recovered purified HIV-1 envelope protein material is agglomerated or assembled into particles of about  $\geq$  2,000,000 molecular weight.
- 25 30. A method in accordance with Claim 24 wherein said HIV-1 envelope protein material is gp160.
31. A method in accordance with Claim 24 wherein said HIV-1 envelope protein material is gp160 which is agglomerated or assembled into particles having a molecular weight of about  $\geq$  2,000,000.
32. Undenatured or native HIV-1 envelope protein material in accordance with Claim 1 in finely divided particle-form and having a molecular weight of about  $\geq$  2,000,000.
- 30 33. Undenatured or native HIV-1 envelope protein material in accordance with Claim 32 wherein said protein material is gp160.
34. Undenatured or native HIV-1 envelope protein material in accordance with Claim 1 having a substantially spherical particle-form in the range about 30-100nM.
- 35 35. Undenatured or native HIV-1 envelope protein material in accordance with Claim 1 wherein said protein material is gp160.
36. An immunogen composition comprising undenatured or native HIV-1 envelope protein material complexed with alum wherein said protein material is present therein at a concentration of about 40 $\mu$ g/ml of said composition.
- 40 37. A method of complexing HIV-1 envelope protein material with alum wherein said protein is complexed with said alum at a pH of about 7.0-7.2.
38. A method in accordance with Claim 37 wherein said HIV-1 envelope protein material is gp160.

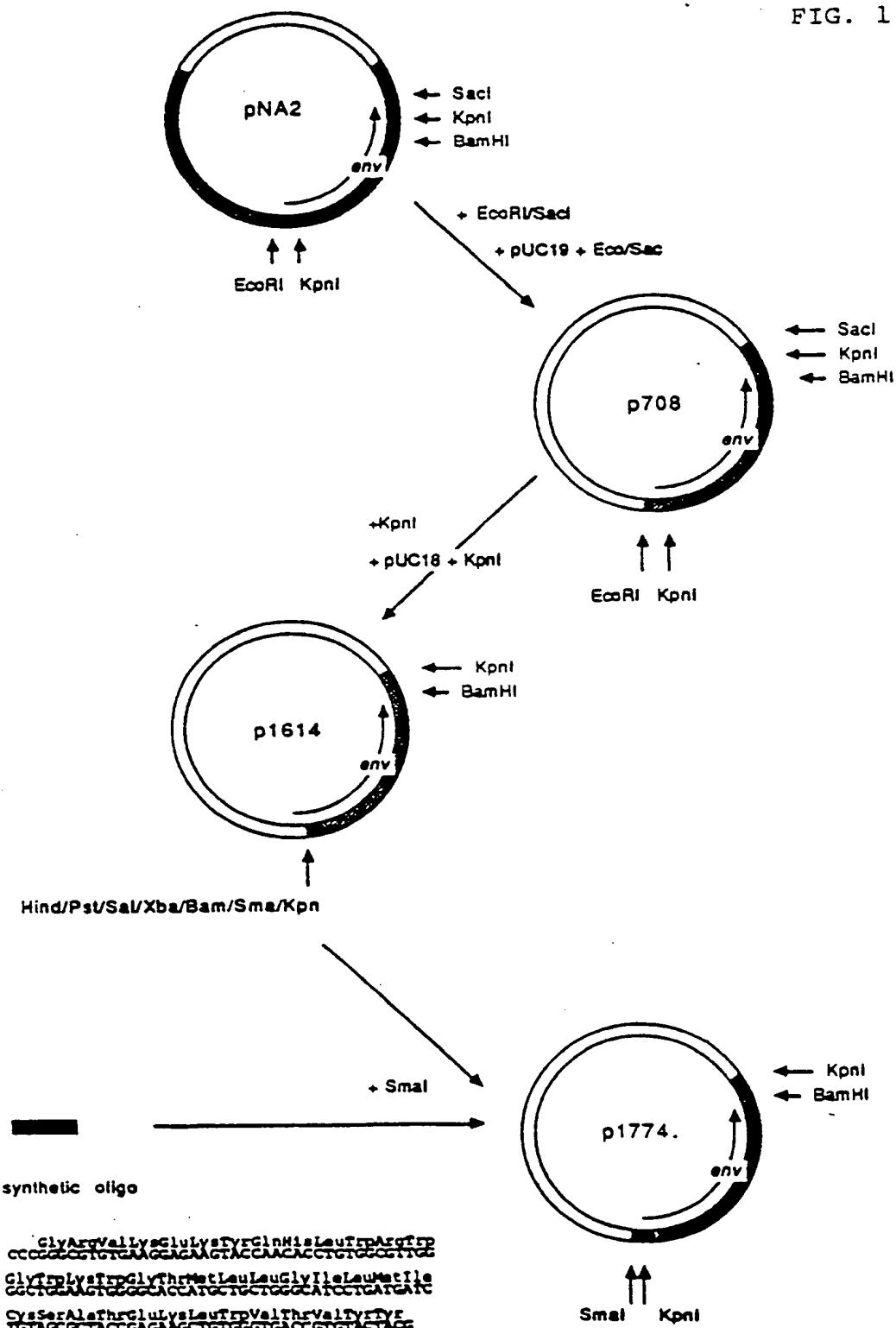
46

50

55

## ISOLATION AND ENGINEERING OF HIV-1 env GENE

FIG. 1



## CONSTRUCTION OF RECOMBINATION VECTOR p3046

FIG. 2

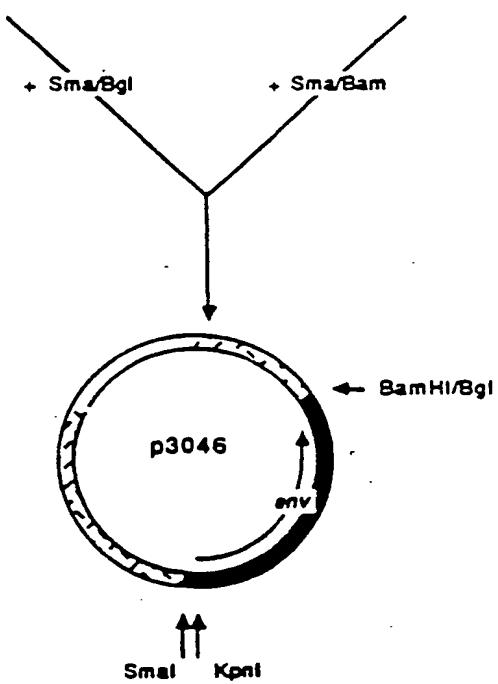
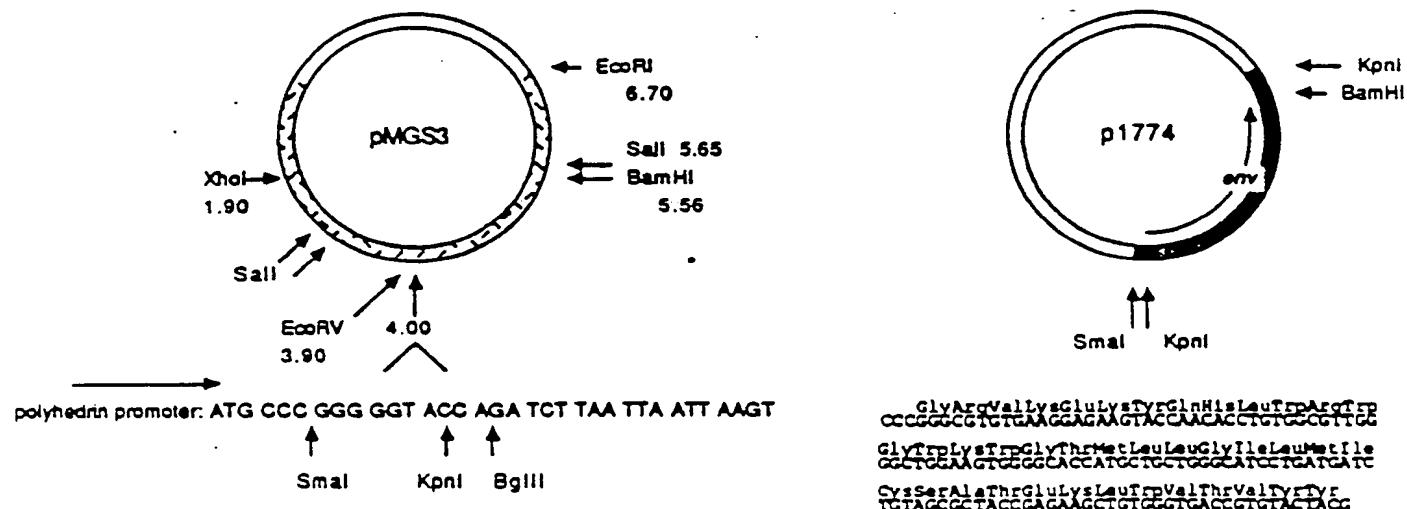


FIG. 3

NUCLEOTIDE SEQUENCE OF DNA FLANKING THE Ac3046 gp160 CODING SEQUENCES

TGCTGATATC ATGGAGATAA TTAAATGAT AACCATCTCG CAAATAAATA  
-100

AGTATTTAC TGTTTCGTA ACAGTTTGT AATAAAAAAA CCTATAATA  
-50

ATG -----/3046/----->TAATTAATTAA GT ACC GAC TCT GGT GAA GAG  
+1 +2253

GAG GAA ATT CTC CTT GAA GTT TCC CTG GTG TTC AAA GTA AAG GAG  
+2287

FIG. 4a

NUCLEOTIDE SEQUENCE AND PREDICTED AMINO ACID SEQUENCE OF  
3046 OPEN READING FRAME

### With enzymes:

	SS						
	AHNNccSS	R		B	B	N	
	vpccrrem	S		b	a	1	
	aaiiFFca	a		v	n	a	
	12111111	1		l	1	4	
	//////						
1	ATGCCCGGGCGTGTGAAGGAGAAGTACCAACACCTGTGGCGTTGGGCTGGAAGTGGGC						
a:	TACGGGCCCCCACACTTCCTCTTCATGGTTGTGGACACCGCAACCCGACCTTCACCCCG						
	MetProGlyArgValLysGluLysTyrGlnHisLeuTrpArgTrpGlyTrpLysTrpGly -						
	B F B	SS F		B			
	sF Nn s	af nP HH		B A			
	po lu t	ua us ha		b l			
	1k a4 X	3N 4t ae		v u			
	21 3H 1	A1 H1 12		1 1			
	/ /			/			
61	ACCATGCTGCTGGGCATCCTGATGATCTGCAGCGCTACCGAGAAGCTGTGGGTGACCGTG						
a:	TGGTACGACGACCCGTAGGACTACTAGACGTCGCGATGGCTCTCGACACCCACTGGCAC						
	ThrMetLeuLeuGlyIleLeuMetIleCysSerAlaThrGluLysLeuTrpValThrVal -						
	R H	B NRK		S			
	s p	a lsp		f			
	a h	n aan		a			
	l l	1 411		N			
	/			l			
121	TACTACGGGGTACCTGTGTGGAAAGGAAGCAACCACACTCTATTGTGCATCAGATGCT						
a:	ATGATGCCCGACACACCTTCCTCGTTGGTGGTGGATAAAACACGTAGTCTACGA						
	TyrTyrGlyValProValTrpLysGluAlaThrThrLeuPheCysAlaSerAspAla -						

FIG. 4b

N	M	T	S	N
t	R	a	H	Ns
d	n	u	a	lp
e	l	g	e	aH
1	1	2	3	31
			6 3	1

AAAGCATATGATAACAGAGGTACATAATGTTGGGCCACACATGCCTGTGTACCCACAGAC  
 181 -----+-----+-----+-----+-----+-----+-----+-----+-----+ 240  
 TTTCGTATACTATGTCTCCATGTATTACAAACCCGGTGTGTACGGACACATGGGTGTCTG  
 a: LysAlaTyrAspThrGluValHisAsnValTrpAlaThrHisAlaCysValProThrAsp -

		N	
M	A	Ns	
a	f	lp	
e	l	aH	
3	3	31	
	/		

CCCAACCCACAAGAAGTAGTATTGGTAAATGTGACAGAAAATTTAACATGTGGAAAAAT  
 241 -----+-----+-----+-----+-----+-----+-----+-----+-----+ 300  
 GGGTTGGGTGTTCTTCATCATAACCATTACACTGTCTTTAAAATTGTACACCTTTTA  
 a: ProAsnProGlnGluValValLeuValAsnValThrGluAsnPheAsnMetTrpLysAsn -

S			S	
fN	M	N N	a	B
a1	n	s l	u	i
Na	l	i a	3	n
13	1	1 3	A	1

N	D	D	A	BH	M
1	r	r	p	sg	
a	a	a	L	1A	o
3	3	1	1	21	2
			/		

FIG. 4c

AATAACCAATAGTAGTAGCGGGGAGAATGATAATGGAGAAAGGAGAGATAAAAAACTGCTCT  
 421 -----+-----+-----+-----+-----+-----+-----+-----+-----+ 480  
 a: TTATGGTTATCATCATGCCCTCTTACTATTACCTCTTCCTCTCTATTTTGACGAGA  
     AsnThrAsnSerSerSerGlyArgMetIleMetGluLysGlyGluIleLysAsnCysSer -  
  
 T  
 t  
 h  
 3  
 2  
                         B N  
                         s s  
                         m i  
                         l l  
  
 TTCAATATCAGCACACAAGCATAAGAGATAAGGTGCAGAAAGAATATGCATTCTTTATAAA  
 481 -----+-----+-----+-----+-----+-----+-----+-----+ 540  
 a: AAGTTATAGTCGTGTTCGTATTCTCTATTCCACGTCTTCTTACGTAAAGAAAATATTT  
     PheAsnIleSerThrSerIleArgAspLysValGlnLysGluTyrAlaPhePheTyrLys -  
  
 R                        A                        M                        D  
 s                        l                        a                        d  
 a                        u                        e                        e  
 l                        l                        3                        1  
  
 CTTGATATAGTACCAATAGATAATACCAGCTATAGGTGATAAGTTGTAACACCTCAGTC  
 541 -----+-----+-----+-----+-----+-----+-----+-----+ 600  
 a: GAACTATATCATGGTTATCTATTATGGTCGATATCCAACIATTCAACATTGGAGTCAG  
     LeuAspIleValProIleAspAsnThrSerTyrArgLeuIleSerCysAsnThrSerVal -  
  
 M                        HHS  B  
 n                        aat  SAsHN  
 l                        eeu  evppc  
 l                        131  calai  
 //  11221  
 //  ///  
 ATTACACAGGCCGTCCAAAGGTATCCTTGAGCCAATTCCCACATACATTATTGTGCCCG  
 601 -----+-----+-----+-----+-----+-----+-----+-----+ 660  
 a: TAATGTGTCCGGACAGGTTCCATAGGAAACTCGGTTAAGGGTATGTAATAAACACGGGGC  
     IleThrGlnAlaCysProLysValSerPheGluProIleProIleHisTyrCysAlaPro -  
  
 SS                        H  S  
 NccSS                    i  M                        Aa                NR  
 crrem                    n  a                    - vu            ls  
 IFFca                   f  e                    - a9            aa  
 11111                  1  2                    26            31  
 //      //  
 GGTGGGTTTGGGATTCTAAATGTAATAATAAGACGTTCAATGGAACAGGACCCATGTACA  
 661 -----+-----+-----+-----+-----+-----+-----+-----+ 720  
 a: CCACCAAAACGCTAAGATTTACATTATTCTGCAAGTTACCTTGTCCIGGEACATGT  
     GlyGlyPheAlaIleLeuLysCysAsnAsnLysThrPheAsnGlyThrGlyProCysThr -

FIG. 4d

		H	
R	R	N1	HH.
s	s	ln	aa
a	a	af	ee
1	1	31	13
		/	

721 AATGTCAGCACAGTACAATGTACACATGGAATCAGGCCAGTAGTATCAACTCAACTGCTG  
 TTACAGTCGTGTCATGTTACATGTGTACCTTAGTCCGGTCATCATAGTTGAGTTGACGAC  
 a: AsnValSerThrValGlnCysThrHisGlyIleArgProValValSerThrGlnLeuLeu -

	S	
M	M	BaX
a	b	guh
e	o	13o
1	2	2A2
		//

781 TTAAATGGCAGTCTAGCAGAAGAAGATGTAGTAATTAGATCTGCCAATTCACAGACAAT  
 AATTACCGTCAGATCGTCTTCTTACATCATTAATCTAGACGGTTAAAGTGTCTGTTA  
 a: LeuAsnGlySerLeuAlaGluGluAspValValIleArgSerAlaAsnPheThrAspAsn -

	N	
R	AsP	R
s	lpv	s
a	uBu	a
1	122	1
	//	

841 GCTAAAACCATAATAGTACAGCTAACACATCTGTAGAAATTAAATTGTACAAGACCCAAC  
 CGATTTGGTATTATCATGTCGACTTGTGTAGACATCTTAATTAAACATGTTCTGGGTTG  
 a: AlaLysThrIleIleValGlnLeuAsnThrSerValGluIleAsnCysThrArgProAsn -

	SBS	
M	AFNascS	M
n	vilutre	a
l	ana9NFC	e
1	2146111	3
	/// //	

901 AACAAATACAAGAAAAAGTATCCGTATCCAGAGGGGACCAAGGGAGAGCATTGTTACAATA  
 TTGTTATGTTCTTTCTAGGCATAGGTCTCCCCTGGTCCCTCTCGTAAACAAATGTTAT  
 a: AsnAsnThrArgLysSerIleArgIleGlnArgGlyProGlyArgAlaPheValThrIle -

FIG. 4e

	T		B
M	c		
a	h		s
e	3		m
3	2		1

GGAAAAAAATAGGAAATATGAGACAAGCACATTGTAAACATTAGTAGAGCAAAATGGAATGCC  
 961 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 1020  
 CCTTTTATCCTTATACTCTGTTCGTGTAAACATTGTAATCATCTCGTTTACCTTACGG  
 a: GlyLysIleGlyAsnMetArgGlnAlaHisCysAsnIleSerArgAlaLysTrpAsnAla -

D	AMN
r	lah
a	uee
1	111
	/

ACTTTAAAACAGATAGCTAGCAAATTAAAGAGAACAAATTGGAAATAATAAAACAATAATC  
 1021 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 1080  
 TGAAATTTGTCTATCGATCGTTAATTCTCTTGTAAACCTTATTATTTGTTATTAG  
 a: ThrLeuLysGlnIleAlaSerLysLeuArgGluGlnPheGlyAsnAsnLysThrIleIle -

PS				
M	DM	ADFMNNpa	M	M
n	ds	vrinlluu	a	n
l	et	aanlaaM9	e	l
1	12	22114416	3	1
	/	////////		

TTTAAGCAATCCTCAGGAGGGGACCCAGAAATTGTAACGCACAGTTTAATTGTGGAGGG  
 1081 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 1140  
 AAATTCGTTAGGAGTCCTCCCTGGGTCTTAAACATTGCGTGTCAAAATTAAACACCTCCC  
 a: PheLysGlnSerSerGlyGlyAspProGluIleValThrHisSerPheAsnCysGlyGly -

RS	RS
sc	sc
aa	aa
11	11
/	/

GAATTTTCTACTGTAATTCAACACAACTGTTAATAGTACTTGTTAAATAGTACTTGG  
 1141 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 1200  
 CTTAAAAAGATGACATTAAAGTTGTGTGACAAATTATCATGAACCAAATTATCATGAACC  
 a: GluPhePheTyrCysAsnSerThrGlnLeuPheAsnSerThrTrpPheAsnSerThrTrp -

FIG. 4f

		E		E		
RS		C	M	C	N	
SC		O	a	O	l	
aa		5	e	5	a	
11		7	3	7	3	
/						
1201	AGTACTGAAGGGTCAAATAACACTGAAGGAAGTGACACAATCACACCTCCOCATGCAGAATA					
a:	TCATGACTTCCCAGTTATTGTGACTTCCTTCACTGTGTTAGTGTGAGGGTACGTCTTAT				1260	
	SerThrGluGlySerAsnAsnThrGluGlySerAspThrIleThrLeuProCysArgIle				-	
		N				
	A	Ns			M	
	f	lp			n	
	1	aH			l	
	3	31			1	
	/					
1261	AAACAATTATAAACATGTGGCAGGAAGTAGGAAAAGCAATGTATGCCCTCCCATCAGT					
a:	TTTGTAAATATTTGTACACCGTCTTCATCCTTTCGTTACATACGGGGAGGGTAGTCA				1320	
	LysGlnPheIleAsnMetTrpGlnGluValGlyLysAlaMetTyrAlaProProIleSer				-	
		F				
	B	S	n			
	b	s	u			
	v	p	4			
	1	1	H			
1321	GGACAAATTAGATGTTCATCAAATATTACTGGGCTGCTATTACAAGAGATGGTGGTAAT					
a:	CCTGTTAACATACAAGTAGTTATAATGACCCGACGATAATTGTTCTCTACCACCATTA				1380	
	GlyGlnIleArgCysSerSerAsnIleThrGlyLeuLeuLeuThrArgAspGlyGlyAsn				-	
	E	S	S	BS		
	C	AMNa	BaX	M M scG	M	F
	O	vblu	guh	n n trs	n	i
	5	aoa9	13o	1 1 NFu	1	n
	7	2246	2A2	1 1 111	1	1
	//	//		/		
1381	AAACAACATGGGTCCGAGATCTTCAGACCTGGAGGAGGOGATATGAGGGACAATTGGAGA					
a:	TTGTTGTTACCCAGGCTCTAGAAGTCTGGACCTCCGCTATACCOCTGTTAACCTCT					1440
	AsnAsnAsnGlySerGluIlePheArgProGlyGlyAspMetArgAspAsnTrpArg					-

FIG. 4g

SS  
 et  
 cy  
 11  
 /  
 AGTGAATTATATAAATATAAAAGTAGTAAAAATTGAACCATTAGGAGTAGCACCCACCAAG  
 1441 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 1500  
 a: TCACTTAATATATTTAATTCATCATTAACTGGTAATCCTCATCGTGGGTGGTTC  
 SerGluLeuTyrLysTyrLysValValLysIleGluProLeuGlyValAlaProThrLys -  
  
 M A SS  
 b l et  
 o u cy  
 2 l 11  
 /  
 GCAAAGAGAAGAGTGGTGCAGAGAGAAAAAGAGCAGTGGGAATAGGAGCTTGTTCTTCTT  
 1501 -----+-----+-----+-----+-----+-----+-----+-----+-----+ 1560  
 a: CGTTCTCTCTCACCAACGTCTCTTTCTCGTACCCCTATCCTCGAAACAAGGAA  
 AlaLysArgArgValValGlnArgGluLysArgAlaValGlyIleGlyAlaLeuPheLeu -  
  
 F F  
 n BH Hn B R  
 u bg hu .. b g s  
 4 va a4 v a a  
 H 11 1H 1 1 1  
 GGGTTCTGGGAGCAGCAGGAAGCACTATGGCGCAGCGTCAATGACGCTGACGGTACAG  
 1561 -----+-----+-----+-----+-----+-----+-----+-----+-----+ 1620  
 a: CCCAAGAACCCCTCGTCGTCTTCGTGATAACCCCGCGTCGAGTTACTGCGACTGOCATGTC  
 GlyPheLeuGlyAlaAlaGlySerThrMetGlyAlaAlaSerMetThrLeuThrValGln -  
  
 F F  
 HH n n BM B D  
 aa u u bn b d  
 ee 4 4 v1 v e  
 13 H H 11 1 1  
 /  
 GCCAGACAATTATTGTCTGATATAGTGCAGCAGCAGAACAAATTGCTGAGGGCTATTGAG  
 1621 -----+-----+-----+-----+-----+-----+-----+-----+-----+ 1680  
 a: CGGTCTGTTAATAACAGACTATATCACGTCTCGTCTTGTAAACGACTCCCGATAACTC  
 AlaArgGlnLeuLeuSerAspIleValGlnGlnGlnAsnAsnLeuLeuArgAlaIleGlu -

FIG. 4h

S	S	BS	TH	BS
f	f	Gsc	ti	sc
a	a	str	hn	tr
N	N	uNF	3f	NF
1	1	111	21	11
		/		/

CAACAGCATCTGTTGCAATCTACAGTCTGGGGCATCAAACAACTCCAGGCAAGAATCCTG  
 1681 -----+-----+-----+-----+-----+-----+-----+ 1740  
 GTTGTGTTAGACAACGTTAGATGTCAGACCCGTAGTTGTTGAGGTCGGTTCTTAGGAC  
 a: GlnGlnHisLeuLeuGlnSerThrValTrpGlyIleLysGlnLeuGlnAlaArgIleL... -

S	BS
a	BA Ssc
u	11 etr
3	nu CNF
A	11 111
	/

GCTGTGGAAAGATACTAAAGGATCAACAGCTCCTGGGGATTGGGGTGCTCTGGAAAA  
 1741 -----+-----+-----+-----+-----+-----+-----+ 1800  
 CGACACCTTTCTATGGATTCTCTAGTTGTCGAGGACCCCTAAACCCCAACGAGACCTTT  
 a: AlaValGluArgTyrLeuLysAspGlnGlnLeuLeuGlyIleTrpGlyCysSerGlyLys -

M	SS	MB
m	et	as
e	cy	em
1	11	11
	/	

CTCATTGCACCACTGCTGCTTGGAAATGCTAGTTGC GTAATAAAATCTCTGGAACAG  
 1801 -----+-----+-----+-----+-----+-----+-----+ 1860  
 GAGTAAACGTGGTGACGACACGGAACCTTACGATCAACCTCATTATTTAGAGACCTTGTC  
 a: LeuIleCysThrThrAlaValProTrpAsnAlaSerTrpSerAsnLysSerLeuGluGln -

BS		H
N sc	F F	i A
l tr	i o	n l
a NF	n k	d u
3 11	1 1	3 1
	/	

ATTTGGAAATAACATGACCTGGATGGAGTGGGACAGAGAAATTAAACATTACACAAGCTTA  
 1861 -----+-----+-----+-----+-----+-----+-----+ 1920  
 TAAACCTTATTGACTGGACCTACCTCACCTGTCTTTAATTGTTAATGTGTTGAAAT  
 a: IleTrpAsnAsnMetThrTrpMetGluTrpAspArgGluIleAsnAsnTyrThrSerLeu -

FIG. 4i

	H	
i		M
n		b
f		c
l		o
	2	

ATACACTCCTAATTGAAGAACGCAAAACCAGCAAGAAAAGAACAGAAATTATTG  
 1921 -----+-----+-----+-----+-----+-----+-----+-----+-----+ 1980  
 a: TATGTGAGGAATTAACTTCTTAGCGTTGGTCGTTCTTACTTGTTCTTAATAAC  
     IleHisSerLeuIleGluGluSerGlnAsnGlnGlnGluLysAsnGlnGluGlnGluLeuLeu -

	GAATTAGATAAATGGGCAAGTTGTGGAATTGGTTAACATAACAAATTGGCTGTGGTAT 1981 -----+-----+-----+-----+-----+-----+-----+-----+-----+ 2040 a: CTTAACATCTATTACCCGTTCAAACACCTTAACCAAATTGTATTGTTAACCGACACCATA GluLeuAspLysTrpAlaSerLeuTrpAsnTrpPheAsnIleThrAsnTrpLeuTrpTyr -
--	---

	M		R
n			s
l			a
1			1

ATAAAATTATTCTATAATGATAGTAGGAGGCTTGGTAGGTTAACAGAACAGATTTGCTGTA  
 2041 -----+-----+-----+-----+-----+-----+-----+-----+-----+ 2100  
 a: TATTTTAATAAGTATTACTATCATCCTCCGAACCATCCAAATTCTTATCAAAACGACAT  
     IleLysLeuPheIleMetIleValGlyGlyLeuValGlyLeuArgIleValPheAlaVal -

	H	
P		
h		
1		

CTTTCTATAGTGAATAGAGTTAGGCAGGGATATTCAACCATTATCGTTTCAGACCCACCTC  
 2101 -----+-----+-----+-----+-----+-----+-----+-----+-----+ 2160  
 a: GAAAGATATCACTTATCTCAATCCGTCCCTATAAGTGGTAATAGCAAAGTCTGGGTGGAG  
     LeuSerIleValAsnArgValArgGlnGlyTyrSerProLeuSerPheGlnThrHisLeu -

	PS	S		
M	AMS	ADFNNpa	aR	M M
n	vne	vrilluu	ua	b b
l	alc	aanaaM9	9e	c c
1	111	2214416	63	2 2
	/			
2161	CCAATCCCAGGGGACCCGACAGGCCCGAAGGAATAGAAGAACAGGTGGAGAGAGAC			2220
a:	GGTTAGGGCTCCCTGGCTGTCCGGCTTCCTTATCTTCTTCCACCTCTCTCTG			
	ProIleProArgGlyProAspArgProGluGlyIleGluGluGluGlyGluArgAsp	-		

FIG. 4j

	S			S		
	B	aX	T	aX	B	
i	uh		a	uh	1	
n	3o		q	3o	n	
1	A2		1	A2	1	
	/			/		

AGAGACAGATCCATTGATTAGTAGTGAACGGATC**T**TAATTAAATTAA  
 2221 -----+-----+-----+-----+--- 2264  
 TCTCTGTCTAGGTAAGCTAATCACTGCCTAGA**T**TAATTAAATT  
 a: ArgAspArgSerIleArgLeuValAsnGlySerEndLeuIleLys-

Enzymes that do cut:

Af13	AluI	ApaLI	Aval	Ava2	BanI	EbvI	Bgl2	BinI	BsmI	Bsp12
BstE2	BstN1	BstX1	DdeI	DraI	Dra2	Dra3	Eco57	FinI	Fnu4R	FokI
Gsul	HaeI	Hae2	Hae3	HgaI	HgiAI	Hhal	Hind3	HinfI	Hpa2	HphI
KpnI	Mael	Mae2	Mae3	Mbo2	Mmel	MnlI	Mst2	NciI	NdeI	NheI
Nla3	Nla4	Nsil	NspB2	NspH1	PpuM1	PstI	Pvu2	RsaI	Sau3A	Sau96
Scal	ScrF1	Secl	SfaN1	SmaI	SspI	StuI	StyI	TaqI	Tth32	Xba2

Enzymes that do not cut:

Aat2	AccI	Af12	Aha2	Apal	Asu2	Avr2	BalI	BamH1	Ban2	Bbv2
Bcl1	BglI	BspH1	BspM1	BspM2	BssH2	Cfr1	Cfr10	ClaI	Dsal	Eco31
EcoB	EcoK	EcoR1	EcoRV	Esp1	Fsp1	Gdi2	HgiE2	Hinc2	HpaI	MluI
Nael	NarI	NcoI	NotI	NruI	PflMI	PmaC1	PvuI	Rsr2	SacI	Sac2
Sall	SfiI	SnaB1	SpeI	SphI	SpiI	ThaI	Tth31	XbaI	XhoI	Xma3
XmnI										

FIG. 4k

NUMBER OF OPEN READING FRAME BASES: 2253

NUMBER OF AMINO ACID CODONS: - 2253 + 3 - 751

Amino Acid	Number	Weight.	Totals
GLY -	53	75.1	3980.3
GLU -	41	147.1	6031.1
ASP -	25	133.1	3327.5
VAL -	48	117.1	5620.8
ALA -	37	89.1	3296.7
ARG -	39	174.1	6793.8
SER -	28	105.1	2942.8
LYS -	42	146.2	6140.4
ASN -	58	132.1	7661.8
MET -	17	149.2	2536.4
ILE -	57	131.2	7478.4
THR -	53	119.1	6312.3
TRP -	26	204.2	5309.2
CYS -	21	121.2	2545.2
TYR -	16	181.2	2899.2
LEU -	61	131.2	8003.2
PHE -	25	165.2	4130.0
SER -	26	105.1	2732.6
GLN -	38	146.2	5555.6
HIS -	11	155.2	1707.2
PRO -	29	115.1	3337.9
<b>TOTALS:</b>	<b>751</b>	<b>98,342.4</b>	
		- H <sub>2</sub> O (751 x 18)	
Total estimated weight of non-glycosylated polypeptide		- 84,824.4	
Total Number of glycosylation sites: 28		x 2100 (wt per oligo saccharide)	
Total Estimated Mol. Wt. of gp160		= 84,824.4 + 58800	
		<u>= 143,624.</u>	

FIG. 5a

## COMPARISON OF LAV-1 AND RECOMBINANT Ac3046 gp160\* SEQUENCE

The sequence and corresponding codons on the top lines are those predicted from the engineering and by Wain-Hobson et al. (1985). The sequence along the bottom of each line is that which was determined for Ac3046 from recombinant viral DNA.

2	Met Pro Gly Arg Val Lys Glu Lys Tyr Gln His ATG CCC GGG CGT GTG AAG GAG AAG TAC CAA CAC 	6250
10	<----- Signal Peptide -----> Leu Trp Arg Trp Gly Trp Lys Trp Gly Thr Met Leu Leu Gly Ile Leu Met Ile Cys Ser CTG TGG CGT TGG GGC TGG AAG TGG GGC ACC ATG CTG CTG GGC ATC CTG ATG ATC TGT AGC 	6310
--\ /----- Extracellular Region (gp120) ----->	Ala Thr Glu Lys Leu Trp Val Thr Val Tyr Tyr Gly Val Pro Val Trp Lys Glu Ala Thr GCT ACC GAG AAG CTG TGG GTG ACC GTG TAC TAC <u>GGG_GTA_CCT</u> GTG TGG AAG GAA GCA ACC 	6370
50	Thr Thr Leu Phe Cys Ala Ser Asp Ala Lys Ala Tyr Asp Thr Glu Val His Asn Val Trp ACC ACT CTA TTT TGT GCA TCA GAT GCT AAA GCA TAT GAT ACA GAG GTC CAT AAT GTT TGG 	6430
70	Ala Thr His Ala Cys Val Pro Thr Asp Pro Asn Pro Gln Glu Val Val Leu Val Asn Val GCC ACA CAT GCC TGT GTA CCC ACA GAC CCC AAC CCA CAA GAA GTC GTC TTG GTC AAT GTG 	6490
90	Thr Glu Asn Phe Asn Met Trp Lys Asn Asp Met Val Glu Gln Met Asp Glu Asp Ile Ile ACA GAA AAT TTT AAC ATG TGG AAA AAT GAC ATG GTC GAA CAG ATG GAT GAG GAT ATA ATC 	6550
110	Ser Leu Trp Asp Gln Ser Leu Lys Pro Cys Val Lys Leu Thr Pro Leu Cys Val Ser Leu AGT TTA TGG GAT CAA AGC CTA AAG CCA TGT GTC AAA TTA ACC CCA CTC TGT GTT AGT <u>TTA</u> 	6610
130	Lys Cys Thr Asp Leu Gly Asn Ala Ser Asn Thr Asn Ser Thr Asn Thr Asn Ser Ser Ser AAG TGC ACT GAT TTG GGG AAT GCT AGT AAT ACC AAT AGT ACT AAT ACC AAT AGT AGT AGC 	6670
150	Gly Glu Met Met Met Glu Lys Gly Glu Ile Lys Asn Cys Ser Phe Asn Ile Ser Thr Ser GGG GAA ATG ATG ATG GAG AAA GGA GAG ATA AAA AAC TGC TCT TTC AAT ATC AGC ACA AGC Arg Ile	6730

FIG. 5b

170 Ile Arg Gly Lys Val Gln Lys Glu Tyr Ala Phe Phe Tyr Lys Leu Asp Ile Ile Pro Ile  
 ATA AGA GGT AAG GTG CAG AAA GAA TAT GCA TTT TTT TAT AAA CTT GAT ATA ATA CCA ATA  
 ||| ||| |A| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| G|| ||| |||  
 ASP Phe Val  
 190 Y Asp Asn Asp Thr Thr Ser Tyr Thr Leu Thr Ser Cys Asn Thr Ser Val Ile Thr Gln Ala  
 GAT AAT GAT ACT ACC AGC TAT ACG TTG ACA AGT TGT AAC ACC TCA GTC ATT ACA CAG GCC  
 ||| ||| (-----) ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
 Arg Ile  
 210 Cys Pro Lys Val Ser Phe Glu Pro Ile Pro Ile His Tyr Cys Ala Pro Ala Gly Phe Ala  
 TGT CCA AAG GTA TCC TTT GAG CCA ATT CCC ATA CAT TAT TGT GGC CCG GCT GGT TTT GCG  
 ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
 Gly  
 230 Y Y Y Ile Leu Lys Cys Asn Asn Lys Thr Phe Asn Gly Thr Gly Pro Cys Thr Asn Val Ser Thr  
 ATT CTA AAA TGT AAT AAT AAG ACG TTC AAT GGA ACA GGA CCA TGT ACA AAT GTC AGC ACA  
 ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
 250 Val Gln Cys Thr His Gly Ile Arg Pro Val Val Ser Thr Gln Leu Leu Leu Asn Gly Ser  
 GTA CAA TGT ACA CAT GGA ATT AGG CCA GTA TCA ACT CAA CTG CTG TTG AAT GGC AGT  
 ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
 Ile Leu  
 270 Leu Ala Glu Glu Glu Val Val Ile Arg Ser Ala Asn Phe Thr Asp Asn Ala Lys Thr Ile  
 CTA GCA GAA GAA GAG GTA GTA ATT AGA TCT GCC AAT TTC ACA GAC AAT GCT AAA ACC ATA  
 ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
 Asp  
 290 Y Y Y Ile Val Gln Leu Asn Gln Ser Val Glu Ile Asn Cys Thr Arg Pro Asn Asn Asn Thr Arg  
 ATA GTA CAG CTG AAC CAA TCT GTA GAA ATT AAT TGT ACA AGA CCC AAC AAC AAT ACA AGA  
 ||| ||| ||| ||| ||| AC| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
 Thr  
 310 Lys Ser Ile Arg Ile Gln Arg Gly Pro Gly Arg Ala Phe Val Thr Ile Gly Lys Ile Gly  
 AAA AGT ATC CGT ATC CAG AGG GGA CCA GGG AGA GCA TTT GTT ACA ATA GGA AAA ATA GGA  
 ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
 330 Asn Met Arg Gln Ala His Cys Asn Ile Ser Arg Ala Lys Trp Asn Ala Thr Leu Lys Gln  
 AAT ATG AGA CAA GCA CAT TGT AAC ATT AGT AGA GCA AAA TGG AAT GCC ACT TTA AAA CAG  
 ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
 350 Ile Ala Ser Lys Leu Arg Glu Gln Phe Gly Asn Asn Lys Thr Ile Ile Phe Lys Gln Ser  
 ATA GCT AGC AAA TTA AGA GAA CAA TTT GGA AAT AAT AAA ACA ATA ATC TTT AAG CAA TCC  
 ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
 370 Ser Gly Gly Asp Pro Glu Ile Val Thr His Ser Phe Asn Cys Gly Gly Glu Phe Phe Tyr  
 TCA GGA GGG GAC CCA GAA ATT GTA ACG CAC AGT TTT AAT TGT GGA GGG GAA TTT TTC TAC  
 ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

6790 .  
 6850 .  
 6910 .  
 6970 .  
 7030 .  
 7090 .  
 7150 .  
 7210 .  
 7270 .  
 7330 .  
 7390 .

FIG. 5c

390 Y Y  
 Cys Asn Ser Thr Gln Leu Phe Asn Ser Thr Trp Phe Asn Ser Thr Trp Ser Thr Glu Gly  
 TGT AAT TCA ACA CAA CTG TTT AAT AGT ACT TGG TTT AAT AGT ACT TGG AGT ACT GAA GGG  
 |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||  
 410 Y  
 Ser Asn Asn Thr Glu Gly Ser Asp Thr Ile Thr Leu Pro Cys Arg Ile Lys Gln Phe Ile  
 TCA AAT AAC ACT GAA GGA AGT GAC ACA ATC ACA CTC CCA TGC AGA ATA AAA CAA TTT ATA  
 |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||  
 430  
 Asn Met Trp Gln Glu Val Gly Lys Ala Met Tyr Ala Pro Pro Ile Ser Gly Gln Ile Arg  
 AAC ATG TGG CAG GAA GTA GGA AAA GCA ATG · TAT GCC CCT CCC ATC AGC GGA CAA ATT AGA  
 |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||  
 450 Y  
 Cys Ser Ser Asn Ile Thr Gly Leu Leu Leu Thr Arg Asp Gly Gly Asn Asn Asn Asn Gly  
 TGT TCA TCA AAT ATT ACA GGG CTG CTA TTA ACA AGA GAT GGT GGT AAT AAC AAC AAT GGG  
 |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||  
 470  
 Ser Glu Ile Phe Arg Pro Gly Gly Asp Met Arg Asp Asn Trp Arg Ser Glu Leu Tyr  
 TCC GAG ATC TTC AGA CCT GGA GGA GAT ATG AGG GAC AAT TGG AGA AGT GAA TTA TAT  
 |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||  
 490  
 Lys Tyr Lys Val Val Lys Ile Glu Pro Leu Gly Val Ala Pro Thr Lys Ala Lys Arg Arg  
 AAA TAT AAA GTA GTA AAA ATT GAA CCA TTA GGA GTA GCA CCC ACC AAG GCA AAG AGA AGA  
 |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||  
 510 <---- gp120 ----- \ / ----- Transmembrane Region (gp 41) ---->  
 Val Val Gln Arg Glu Lys Arg Ala Val Gly Ile Gly Ala Leu Phe Leu Gly Phe Leu Gly  
 GTG GTG CAG AGA GAA AAA AGA GCA GTG GGA ATA GGA GCT TTG TTC CTT GGG TTC TTG GGA  
 |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||  
 530  
 Ala Ala Gly Ser Thr Met Gly Ala Arg Ser Met Thr Leu Thr Val Gln Ala Arg Gln Leu  
 GCA GCA GGA AGC ACT ATG GGC GCA CGG TCA ATG ACG CTG ACG GTA CAG GCC AGA CAA TTA  
 |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||  
 550  
 Leu Ser Gly Ile Val Gln Gln Gln Asn Asn Leu Leu Arg Ala Ile Glu Ala Gln Gln His  
 TTG TCT GGT ATA GTG CAG CAG CAG AAC AAT TTG CTG AGG GCT ATT GAG GCG CAA CAG CAT  
 |||| |||| |||A| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |(---)| |||| |||| ||||  
 570 Asp  
 Leu Leu Gln Leu Thr Val Trp Gly Ile Lys Gln Leu Gln Ala Arg Ile Leu Ala Val Glu  
 CTG TTG CAA CTC ACA GTC TGG GGC ATC AAG CAG CTC CAG GCA AGA ATC · CTG GCT GTG GAA  
 |||| |||| |||| TCT |||| |||| |||| |||| |||A| |||A| |||| |||| |||| |||| |||| |||| |||| |||| ||||  
 590 Ser Lys Gln  
 Arg Tyr Leu Lys Asp Gln Gln Leu Leu Gly Ile Trp Gly Cys Ser Gly Lys Leu Ile Cys  
 AGA TAC CTA AAG GAT CAA CAG CTC CTG GGG ATT TGG GGT TGC TCT GGA AAA CTC ATT TGC  
 |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||  
 7450  
 7510  
 7570  
 7630  
 7690  
 7750  
 7810  
 7870  
 7930  
 7990  
 8050

FIG. 5d

610                  Y                  Y  
 Thr Thr Ala Val Pro Trp Asn Ala Ser Trp Ser Asn Lys Ser Leu Glu Gln Ile Trp Asn  
 ACC ACT GCT GTG CCT TGG AAT GCT AGT TGG AGT AAT AAA TCT CTG GAA CAG ATT TGG AAT        8110  
 |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||

Y                  Y  
 Asn Met Thr Trp Met Glu Trp Asp Arg Glu Ile Asn Asn Tyr Thr Ser Leu Ile His Ser  
 AAC ATG ACC TGG ATG GAG TGG GAC AGA GAA ATT AAC AAT TAC ACA AGC TTA ATA CAT TCC        8170  
 |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||  
 His

650  
 Leu Ile Glu Glu Ser Gln Asn Gln Gln Glu Lys Asn Glu Gln Glu Leu Leu Glu Leu Asp        8230  
 TTA ATT GAA GAA TCG CAA AAC CAG CAA GAA AAG ATT GAA CAA GAA TTA TTG GAA TTA GAT  
 |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||

670                  Y  
 Lys Trp Ala Ser Leu Trp Asn Trp Phe Asn Ile Thr Asn Trp Leu Trp Tyr Ile Lys Ile        8290  
 AAA TGG GCA AGT TTG TGG AAT TGG TTT AAC ATA ACA AAT TGG CTG TGG TAT ATA AAA ATA  
 |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||  
 Leu

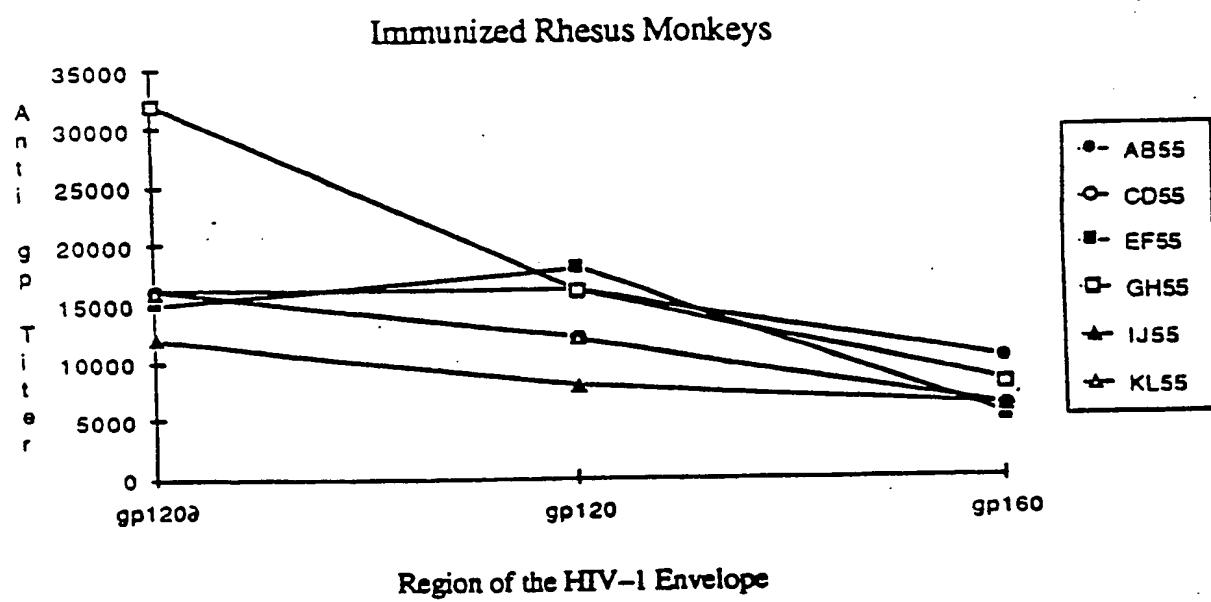
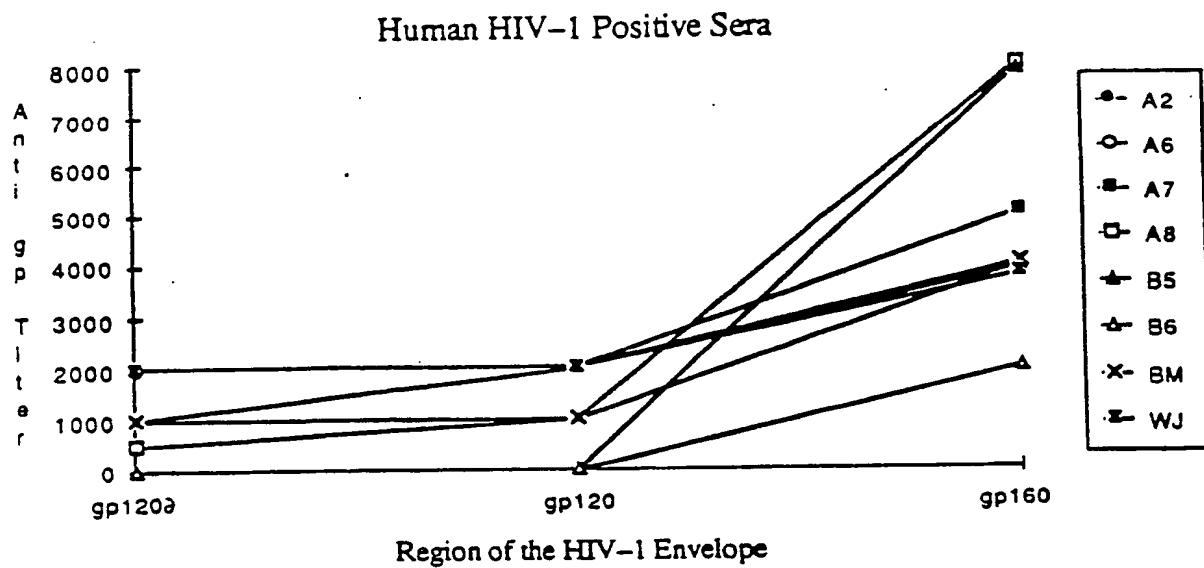
690  
 Phe Ile Met Ile Val Gly Gly Leu Val Gly Leu Arg Ile Val Phe Ala Val Leu Ser Ile        8350  
 TTC ATA ATG ATA GTA GGA GGC TTG GTA GGT TTA AGA ATA GTT TTT GCT GTA CTT TCT ATA  
 |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||

710  
 Val Asn Arg Val Arg Gln Gly Tyr Ser Pro Leu Ser Phe Gln Thr His Leu Pro Thr Pro        8410  
 GTG AAT AGA GTT AGG CAG GGA TAT TCA CCA TTA TCG TTT CAG ACC CAC CTC CCA ACC CCG  
 |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||  
 Ile

730  
 Arg Gly Pro Asp Arg Pro Glu Gly Ile Glu Glu Glu Gly Glu Arg Asp Arg Asp Arg        8470  
 AGG GGA CCC GAC AGG CCC GAA GGA ATA GAA GAA GAA GGT GGA GAG AGA GAC AGA GAC AGA  
 |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||

750                  Y  
 Ser Ile Arg Leu Val Asn Gly Ser END  
 TCC ATT CGA TTA GTG AAC GGA TCT TAA TTA ATT AA

FIG. 6





Europäisches Patentamt  
European Patent Office  
Office européen des brevets

(11) Publication number:

0 327 180  
A3

(12)

## EUROPEAN PATENT APPLICATION

(21) Application number: 89200230.4

(51) Int. Cl.5. A61K 39/21, C07K 15/04,  
C12N 15/49

(22) Date of filing: 02.02.89

(30) Priority: 03.02.88 US 151976

(71) Applicant: MicroGeneSys, Inc.

(43) Date of publication of application:  
09.08.89 Bulletin 89/32

400 Frontage Road  
West Haven Connecticut 06437(US)

(84) Designated Contracting States:  
AT BE CH DE ES FR GB GR IT LI LU NL SE

(72) Inventor: Smith, Gale E.

(88) Date of deferred publication of the search report:  
18.07.90 Bulletin 90/29

125 Michael Drive  
Guilford Connecticut 06437(US)

Inventor: Cochran, Mark A.

415 Brooksvale Avenue  
Hamden, Connecticut 06518(US)

Inventor: Volfovitz, Franklin  
123 York Street

New Haven, Connecticut 06511(US)

(54) Vaccine containing polypeptides derived from the envelope gene of human immunodeficiency virus type 1.

(57) An Acquired Immunodeficiency Syndrome (AIDS) vaccine containing the Human Immunodeficiency Virus, Type-1 (HIV-1) envelope proteins is produced from cloned HIV-1 envelope genes in a baculovirus-insect cell vector system. The recombinant HIV-1 proteins are purified, assembled into particles and then adsorbed on an aluminum phosphate adjuvant. The resulting adsorbed recombinant HIV-1 virus envelope protein formulation is highly immunogenic in animals and elicits antibodies which bind to the HIV-1 virus envelope and neutralize the infectivity of the virus in in vitro tests.

EP 0 327 180 A3



DOCUMENTS CONSIDERED TO BE RELEVANT			CLASSIFICATION OF THE APPLICATION (Int. Cl. 4)
Category	Citation of document with indication, where appropriate, of relevant passages	Relevant to claim	CLASSIFICATION OF THE APPLICATION (Int. Cl. 4)
X	PROCEEDINGS NATIONAL ACADEMY SCIENCES, USA, vol. 84, october 1987, pages 6924-6928; J.R. RUSCHE et al.: "Humoral immune response to the entire human immunodeficiency virus envelope glycoprotein made in insect cells" * Whole document *	1-16, 18-19, 35-38	A 61 K 39/21 C 07 K 15/04 C 12 N 15/00
X	GB-A-2 181 435 (ONCOGEN) * Claims *	1-16, 18-19, 35-38	
X	EP-A-0 243 029 (UNITED STATES OF AMERICA) * Columns 6-7 *	1-13, 35-38	
X	JOURNAL OF VIROLOGICAL METHODS, vol. 18, no. 4, 1987, pages 243-255; T.J. PALKER et al.: "Purification of envelope glycoproteins of human T cell lymphotropic virus type 1 (HTLV-I) by affinity chromatography" * Pages 244-245 *	24-28	
D, P X	EP-A-0 265 785 (MICROGENESYS INC.) * Whole document *	1-21, 35-38	A 61 K C 12 P C 12 N
P, X	CHEMICAL ABSTRACTS, vol. 111, no. 1, 3rd July 1989, page 579, abstract no. 5882j, Columbus, Ohio, US; N. BARRETT et al.: "Large-scale production and purification of a vaccinia recombinant-derived HIV-1 gp160 and analysis of its immunogenicity", & AIDS RES. HUM. RETROVIRUSES 1989, 5(2), 159-71 * Whole abstract *	1-10	
The present search report has been drawn up for all claims			TECHNICAL FIELDS SEARCHED (Int. Cl.4)
Place of search		Date of completion of the search	Examiner
THE HAGUE		02-02-1990	ALVAREZ Y ALVAREZ C.
CATEGORY OF CITED DOCUMENTS			
X : particularly relevant if taken alone Y : particularly relevant if combined with another document of the same category A : technological background O : non-written disclosure P : intermediate document			
T : theory or principle underlying the invention E : earlier patent document, but published on, or after the filing date D : document cited in the application L : document cited for other reasons ..... & : member of the same patent family, corresponding document			